AD-A241 462

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

2

...timated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, ...and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this ...his burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson ...the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

...EPORT DATE

3. REPORT TYPE AND DATES COVERED
FINAL, 01 Jun 90 to 30 Nov 90

CONFERENCE ON VISUAL INFORMATION ASSIMILATION
IN MAN AND MACHINE

5. FUNDING NUMBERS
AFOSR-90-0280
61102F, 2304/A7

6. AUTHOR(S)
Ramesh Jain

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Dept of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI 48109

8. PERFORMING ORGANIZATION REPORT NUMBER
AFOSR-TR- 91 0791

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NM
Bldg 410
Bolling AFB DC 2033206448

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFOSR-90-0280

DTIC
ELECTE
OCT 11 1991
D

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

In June 27-29, the Artificial Intelligence Laboratory of the University of Michigan, in cooperation with the Cognitive Science and Machine Intelligence Laboratory, organized a conference on visual information assimilation. The primary funding agency for the conference was the Air Force Office of Scientific Research. The conference was successful in bringing together a diverse group of participants. About 100 people attended the conference, with cross-disciplinary attendees from both within the University of Michigan and outside.

91-13068

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std 239-18
298-102

91 10 10 ...

Conference on Visual Information Assimilation in Man and Machine
Final Report

Abstract

In June 27-29, 1990, the Artificial Intelligence Laboratory of the University of Michigan, in cooperation with the Cognitive Science and Machine Intelligence Laboratory, organized a conference on visual information assimilation. The primary funding agency for the conference was the Air Force Office of Scientific Research. The conference was succesful in bringing together a diverse group of participants. About 100 people attended the conference, with cross-disciplinary attendees from both within the University of Michigan outside.

Conference on Visual Information Assimilation
in Man and Machine
Final Report

1 Introduction

In June 27-29, 1990, the Artificial Intelligence Laboratory of the University of Michigan, in cooperation with the Cognitive Science and Machine Intelligence Laboratory, organized a conference on visual information assimilation. The primary funding agency for the conference was the Air Force Office of Scientific Research. The conference was organized by graduate students at the University of Michigan: Sarvajit Sinha, Yuval Roth, and Mark Wiesmeyer all from the Department of Electrical Engineering and Computer Science (EECS). Four University of Michigan Professors: Ramesh Jain (EECS), Terry Weymouth (EECS), Laurence Maloney (Psy.), and Stephen Easter (Bio.) supervised the organization of the conference and provided requisite guidance to these graduate students.

2 Goals

The goal of the conference were to

(a) bring together researchers from psychology, neurobiology, and computer science,

(b) discuss emerging themes that will play a key role in our understanding of how both natural and artificial vision systems integrate information from different visual modalities.

Professor Ramesh Jain presented the aims of the conference in his introductory statement. Amongst the questions he asked the speakers to address were:

1. What are the possible representations for each visual modality?

2. How are contradictions in the visual information resolved?

3. Is there (and should computer vision systems have) one coherent representation of the environment?

4. At what point in visual processing does (should) goal or context related information become important and influence processing?

The presentations in the conference attempted to focus on these questions. The talks could be roughly divided into the following areas– synthesis, or creation of algorithms/models of the visual system, mostly from the computer vision community, and analysis of the visual system from the psychology and neurophysiology viewpoints. The analysis area could be further divided into studies concentrating on behavior (the psychology community) versus those concentrating on mechanisms (neurobiology).

The goal of making machines that see is tightly bound to the goal of understanding how natural vision functions. It is true that computer vision algorithms can be written without indepth knowledge of vision in biological system, but it may be useful to look at biological systems for guidance.

Human vision is the best example of a perceptual system that uses the visible spectrum for cognition. The visual system sees and understands what it sees. *If we knew how humans see then* a *(undoubtedly complex) system* could be created to do this task.

The flip side of the coin is that psychology and neuroscience can also use computational vision as a tool to provide their theories of vision with rigor that would otherwise be difficult to prove or justify. Ultimately, the goals of both are to understand the mechanisms of vision, but at different grain sizes and through the use of different tools.

3 Speakers

The following speakers presented their latest research within the theme of the conference:

1. Irving Biederman (Minnesota): Human Object Recognition

2. Patrick Cavanaugh (Montreal): 3-D Representation

3. Daniel Green(Michigan): Control of Visual Sensitivity

4. Allen Jepson (Toronto): Lattice Theory of Perception
5. Stephen Kosslyn (Harvard): Components of High-level Vision
6. Geoffrey Loftus (Washington):Why it's annoying to watch slides with the room lights still on: Degradation of Visual Information Processing
7. Laurence Maloney (NYU): Calibration of Simple Visual Systems by Comparison Across Eye Movements
8. Misha Pavel (Stanford): Integration of Motion Information
9. Whitman Richards (MIT): Lattice Theory of Perception
10. Barry Richmond(NIMH): How Single Neuronal Responses Represent Picture Features Using Multiplexed Temporal Codes
11. Klaus Schulten(Illinois): A Self Organized Network for Feature Extraction
12. John Tsotsos (Toronto): Vision, Complexity and Attention
13. Brian Wandell (Stanford): Color Appearance

A program containing abstracts of each speakers talk is enclosed herewith.

4 Summary

We divide the issues discussed in the meeting into the following topics:

4.1 Representation

The representation of objects is a problem that can be viewed in many different ways. At the signal level, we see the representation in terms of shape identifiers like stereo and motion (Pavel), and feature-indexed modalities such as color (Wandell), texture etc. Higher level representations include geons (Biederman), and symbolic systems with relational descriptions such as on-top-of etc. Cavanagh's thesis is that vision proceeds from a 2-D match which guides a 3-D model construction. Kosslyn summarized the components of high level vision, and a mapping of these onto brain locations-

1. A visual buffer with an attention window
2. An encoding of object properties
3. An encoding of spatial properties
4. A multimodal associative memory
5. Top-down hypothesis testing

At the signal level, Barry Richmond presented the importance of understanding the functioning of a single neuron. His talked showed that how a single neuronal response represents picture features using multiplexed temporal codes, and how this decided how the response was interpreted.

4.2 Computation

The representation of the computations involved in visual processing include neural nets (Biederman, Schulten) and Richards and Jepson's lattice theory. They suggest that perception is the identification of a local maxima in a fault lattice whose nodes contain interpretations consistent with the data and internal models of the world. Maloney presented methods for the calibration of visual systems whose photoreceptors are initially at unknown locations. These algorithms compensate for spherical abberation, and compensate for irregularities in the photoreceptor lattice. Although this is intended as a model of biological system, yet might it not be used as the basis for a computer vision/robotics calibration?

4.3 Problems

Some of the interesting issues to evolve out of the discussion on representation were:

- Invariance of features across different conditions in the environment: This involves not only invariance across illumination conditions (Green, Loftus, Wandell), but also across translation or rotation. Color, for example, is a perceptual code which summarizes our estimate of the surface reflectance of an object. Although it is possible to design a trichromatic system that obtains these estimates under all phases of daylight, research suggests that the human eye does not achieve this goal (Wandell).

- Graceful degradation of the system: This was touched upon by Kosslyn.
- Context: Richards and Jepson suggest that perceptions are interpretations which locally maximize the application of our knowledge about things in the world.
- Attention: The concept of attention(Kosslyn, Tsotsos) received attention, as a mechanism to reduce the computational complexity of the visual task.

5 Conclusion

We feel we succeeded in bringing together a diverse group of participants. About 100 people attended the conference, with cross-disciplinary attendees from both within the University of Michigan, and 40 external participants.

At the end of the conference, an evaluation sheet was passed around. Some of the strengths mentioned by the participants were

- Many distinguished speakers.
- Excellent information content by top-notch speakers.
- Good discussions on several topics of central interest during the presentations as well as through informal interactions during breaks.
- Adequate time apportioned for presentations and discussions every day.

Amongst constructive criticisms we received, we feel the following will be useful for other conferences:

- Needed more presentations and studies of applied research.
- Might try having discussion panels where people could integrate talks with one another.
- Should have video-taped the talks.

As an outcome of the conference, the speakers were invited to submit full-length papers for publication. This has since been turned into an open call-for-papers for a special issue of I.E.E.E. Transactions on Systems, Man and Cybernetics. The due date for papers is April 28th, 1991, with publication in April 1992.